



King's Research Portal

DOI:

[10.1007/978-1-84628-779-4_15](https://doi.org/10.1007/978-1-84628-779-4_15)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Ginzburg, J., & Macura, Z. (2007). Lexical Acquisition with and without Metacommunication. In *Emergence of Communication and Language: Part III* (Vol. III, pp. 287-301). Springer London. https://doi.org/10.1007/978-1-84628-779-4_15

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

15

Lexical Acquisition with and without Metacommunication

Jonathan Ginzburg and Zoran Macura*

15.1. Introduction

A central concern of work on the evolution of language has been to offer an account for the emergence of syntactically complex structure, which underwrites a compositional semantics. In this chapter we consider the emergence of one class of utterances which illustrate that semantic expressiveness is not correlated with syntactic complexity, namely metacommunicative interaction (MCI) utterances. These are utterance acts in which conversationalists acknowledge understanding or request clarification. We offer a simple characterisation of the incremental change required for MCI to emerge from an MCI-less linguistic interaction system. This theoretical setting underpins and motivates the development of an ALife environment in which the lexicon dynamics of populations that possess and lack MCI capabilities are compared.

We ran a series of experiments whose initial state involved agents possessing distinct lexicons and whose end state was one in which all agents associated meanings with each word in a lexicon. The main effect demonstrated, one we dub the *Babel effect*, is that the convergence rate of a population that relies exclusively on introspection is intrinsically bounded and, moreover, this bound decreases with an increasing population. This bound seems to disappear once agents are endowed with clarification requests.

In natural language, semantic expressiveness is not correlated with syntactic complexity. A key feature of natural language, which provides a striking instance of *syntactically underdetermined* semantic complexity, is metacommunicative interaction (MCI)—utterance acts in which conversationalists acknowledge understanding or request clarification. (1b) exemplifies such a syntactically simple form which, nonetheless, in context can acquire a highly complex content:

- (1) (a) A: Did Bo leave?
 (b) B: Bo?; (“Bo?” can mean in this context *Are you asking if Bo, of all people, left* or *Who were you referring to as Bo?*).

*The authors are listed in alphabetical order.

Indeed natural language possesses forms whose sole meaning concerns MCI, as exemplified by (2), a form whose sole use is to query an antecedently uttered polar interrogative whose subject has unclear reference:

(2) Do I like who?

The need to verify that mutual understanding among interlocutors has been achieved with respect to any given utterance—and engage in discussion of a clarification request if this is not the case—is one of the central organizing principles of conversation (Schegloff 1992; Clark 1996). However, hitherto there has been little work on the emergence of MCI meaning. Communicative interaction is fundamental to evolution of grammar work, since it is interactions among communicating agents that leads an initial ‘agrammatical’ system to evolve into a grammar (with possible, concomitant phylogenetic modification; see e.g. (Briscoe 2000; Kirby 2000). However, given an I-language¹ perspective, the communicative aspect as such is not internalized in the grammar (though see (Steels 1998)). Consequently, such models of evolution of grammar cannot explain the existence of forms whose meaning is intrinsically MCI oriented.

In this paper we offer a simple characterization of the incremental change required for MCI to emerge from an MCI-less communicative interaction system. We discuss the evolutionary background in which MCI might arise and become adaptive. Finally, we report on a series of experiments we ran using an ALife environment in which the lexicon dynamics of populations that possess and lack MCI capabilities are compared. These experiments reveal some clear differences in the lexicon dynamics of populations that acquire words solely by introspection contrasted with populations that learn using MCI or using a mixed strategy of introspection and MCI.

15.2. Metacommunicative Interaction and Evolution of Language

15.2.1. *The Significance of MCI for a Linguistic Community*

By metacommunicative interaction one means any interaction that comments about the communicative process underlying an utterance. More specifically, the commonest MCI utterances are: acknowledgements that an utterance has been understood, clarification requests (CRs) in which an unclear aspect of the utterance is queried, and corrections, where indications are provided of

¹ Following Chomsky (as clarified by Hurford), a distinction is sometimes made between ‘I language’ — language as represented in the brains of the population and ‘E-language’ — language that exists as utterances in the arena of use. Ginzburg (2000) dispute the dichotomy particularly given the need for a view of language that accommodates MCI.

15. Lexical Acquisition with/without Metacommunication 289

erroneous assumptions concerning naming, concepts associated with predicates etc. Example (3) below, from the London Lund corpus, contains a CR (utterance (b)), a correction (utterance (d)), and an acknowledgement (utterance (e)):

- (3) A(a): did you also scotch that other story which is something like was he
wasn't he refused the chair in Oxford
a(b): who
A(c): Skeat, wasn't he refused
a(d): that's Meak
A(e): oh Meak, yes
(London Lund S.1.9, p. 245)

What significance does MCI have for linguistic interaction within a community? MCI is redundant in so far as the communication channel, i.e. that which mediates between speaker and addressee, is perfect or close to that. The need for MCI arises when the communication channel is intrinsically liable to breakdown. If natural language resembled formal languages like first order predicate calculus (as often implicitly assumed in evolution of language work, see e.g. (Kirby 2000)), then problems with the communication channel would be restricted to actual physical problems with the speech signal (mishearing, mispronunciation, noise and the like), problems that affect just about any naturally occurring communicative interaction system. However, natural language diverges radically from first order predicate calculus in its context dependence. This manifests itself in (at least) three phenomena:

- (4) (a) *indexicality*: words like 'I', 'you', 'here', 'now', that are resolved relative to the ongoing speech situation.
(b) *anaphoricity*: words and phrases that are resolved relative to semantic values established by previous utterances (e.g. pronouns, non sentential utterances etc.).
(c) *ambiguity*: words and phrases which possess multiple senses, one of which is utilised in a given context.

Moreover, even a language like first order predicate calculus used by agents who can reflect about intentions underlying communicative acts, will give rise to the sort of inferences that have come to be known as Gricean conversational implicatures (Grice 1989). These add an extra layer of uncertainty to the communicative process.

Given this, acknowledgements, CRs and corrections are a key communicative component for a linguistic community. They serve as devices for allaying worries about miscommunication (acknowledgements) or for reducing mismatches about the linguistic system among agents (CRs and corrections). That is, they serve as a device for ensuring a certain state of equilibrium or lack of divergence gets maintained within a linguistic community.

15.2.2. *The Emergence of MCI: Basic Ingredients*

Given the importance that MCI has for linguistic interaction, some fundamental questions that need to be answered are:

- (5) (a) Under what circumstances does a linguistic interaction system without MCI evolve into one that has MCI?
- (b) What mechanisms are involved in such a development?
- (c) Why is the resulting interaction system maintained?

(5a,b) are questions to which we can offer only sketchy suggestions at present (see discussion in the following page); the main issue we will contend with is (5c). In order to address these issues, we need to fix what we mean by *an interaction system with MCI*. In the literature on the semantics and pragmatics of dialogue a number of interaction systems have been defined where in addition to the regular illocutionary acts (assertion, querying, commanding etc.), also additional grounding acts (e.g. acknowledgements) are available (see e.g. (Poesio and Traum 1997)) and also systems where clarification requests are available (see e.g. (Ginzburg and Sag 2000; Ginzburg and Cooper 2004)). Such systems assume that as a preliminary to the processing of an utterance *u* an addressee A checks whether she understands *u*. If she does, A optionally responds with an acknowledgement, and then reacts in the conventional way to the utterance (accepting/disputing an assertion, answering a query, and so on.). On the other hand, if A does not fully understand *u*, A poses a query that requests clarification concerning the unclear aspect of *u* (e.g. inability to resolve a referent, unfamiliarity with or mishearing of a word, etc.) using a number of predefined operations on utterances and utterance meanings.

Poesio and Traum (1997), Ginzburg and Sag(2000), Ginzburg and Cooper (2004) show how existing formal frameworks for grammatical/semantic processing of MCI-less natural language can be extended to process natural language that includes MCI utterances such as acknowledgements and CRs. To understand what is involved, though, one can restrict attention to much simpler systems. We mention two here discussed originally in (Ginzburg 2001).

The utt(erance) ack(nowledgement) Game. In this game, given an utterance u_0 consisting of a string ($\text{word}_1 \dots, \text{word}_i, \dots, \text{word}_n$) by the master, the novice may respond with the utterance u_1 : word_i . In this context, this utterance is assigned content: *novice acknowledges that an utterance including the word word_i happened*. This fact now becomes part of the novice's and master's common ground. What capabilities does playing *utt-ack* game require from the novice?

- Phonological imitation and segmentation module (can be played in one word mode, i.e. game does not require novice to have syntactically complex capabilities)
- Ability to form mutual beliefs

The reward for playing this game is shared interaction with the master. Who can play this game?

15. Lexical Acquisition with/without Metacommunication 291

- 01 – Human neonates: the initial stage of speech consists largely in playing this
- 02 game (Bates 1979; Ninio 1996)
- 03 – Chimps: (Greenfield and Rumbaugh 1993)

04 *A Rudimentary Game with CRs: The **ack-huh?** Game.* Given an utterance u_0 ,

05 the responder may acknowledge the utterance or pose a simple CR querying the

06 content of u_0 ? For instance:

07

08 (6) Master: You want the ball? Novice: (i)huh?/(ii)ball?

09

10 What additional capabilities does playing *ack-huh* require from the novice?

- 11 – Querying
- 12 – The ability to form questions querying the contents of antecedently uttered
- 13 utterances
- 14 – No requirement for syntax

15

16 Who can play this game?

- 17
- 18 – Human neonates (from approx 20 months)
- 19 – **Not** chimps: (Greenfield and Rumbaugh 1993)

20

21 The key feature of these games is at the level of ontology, namely the possibility

22 of reference to utterances and sub-utterances and their properties. In particular,

23 agents capable of playing the *ack-huh?* game require a notion of synonymy

24 between utterances (i.e. the ability to reformulate in a way that preserves content),

25 otherwise any metacommunicative-oriented discussion will be circular. Thus, the

26 simplest agent with the ability to *discuss* a CR is an agent who can communicate

27 contents such as “I don’t understand (previous-utterance)” and “What do you

28 mean (previous-utterance)”. Given an agent who can reflect and form questions

29 about entities in the domain, this means that once ‘say’ and ‘mean’ predicates are

30 in the language, then basic clarification requests can be expressed. Consequently,

31 the emergence of metacommunicative interaction-oriented utterances that go

32 beyond mere acknowledgement, as exemplified in the *ack-huh?* game, can be

33 viewed as an instance of the problem of how vocabulary emerges to talk about

34 a class of entities in a domain, given the need/desire to do so. We speculate that

35 MCI has emerged as an interactional device that keeps members of a linguistic

36 community from diverging too widely from each other’s linguistic capabilities,

37 say in terms of their basic vocabulary.

38 The plausibility of this speculation can be assessed by converting it into more

39 concrete questions such as the following:

- 40 (7) (a) In a community with minor but random lexical differences where some
- 41 people use clarification requests, whereas others do not, do the clarifi-
- 42 cation request users gain an advantage?
- 43 (b) Given a community A where clarification requests do not get expressed,
- 44 and community B where they do, how do the two communities evolve
- 45 with respect to vocabulary drift.

In the following sections we present some results obtained from an environment which simulates simple linguistic interactions with agents who are introspective or use CRs when encountering unknown words.

15.3. An ALife Simulation

15.3.1. *Basic Properties*

The approach we are following, along with many other researchers, employs computational simulations of a population of distributed, autonomous communicative agents endowed with some linguistic capacities. The agents interact via language games, and the outcome can give insight into the particular phenomena that is being investigated.

We are currently running artificial life simulations on a population of agents with dialogue capacities. The model is built using RePast (developed by Collier et al. 2003 and ROAD), a set of Java libraries that allow programmers to build simulation environments. The running of the simulation is divided into time steps or ‘ticks’, and at each tick some action occurs using the results of previous actions as its basis. Agents are created and placed in an environment in which they are able to wander around in search of ‘food’ resources. Agents are endowed with a vision capacity in order to see food resources as well as other agents. Upon meeting, the two agents enter a dialogue by playing a naming game where the speaker chooses a food resource in his field of vision, and sends a representation of it to the hearing agent, which in turn tries to interpret it.

15.3.2. *Agents*

In any multi-agent model the most basic component is the agent. The properties of this agent depend on what we want the agent to do. Since we are modelling a community of communicating agents in a spatial environment, each agent is endowed with the ability to walk, see and communicate. The environment consists of different food resources (i.e. plants) that the agents can see and talk about. An agent walks around the environment in a random fashion and this random probability is the same for every agent. An agent can also perceive other agents and plants that are close by. Agents can make syntactically simple utterances—essentially one consisting of a single word. Each agent’s lexicon stores the ‘meaning—representation’ tuples for the different plants in the environment (e.g. plant-type—plant-word).

Communication is a two sided process involving an intrinsic asymmetry between speaker and addressee: when talking about a plant, the speaking agent necessarily has a lexical representation of the plant, which he sends to the hearing agent. There is no necessity, however, that the addressee agent is able to interpret this utterance. If unable to do so (meaning that the hearing agent doesn’t have the word in her lexicon) the way that the agent tries to ground it depends on the agent’s type.

Two types of communicative agents exist in the model; agents capable of making a clarification request (CR agents) and those incapable of doing so (introspective agents). A CR agent can resort to a clarification request upon hearing an unknown plant-word. The speaking agent answers this clarification request by giving the meaning-representation to the addressee, who is then able to store it in her lexicon.

An introspective agent, on the other hand tries to guess the meaning of an unknown plant-word instead of resorting to a clarification request. Upon hearing an unknown plant-word the agent looks around her and for each plant that she sees she increases the association score of the plant-word with the plants in her field of vision. This is stored in her temporary lexicon (of unknown plant-words) until she has sufficient information to pick a meaning for the unknown word (viz. associate a specific plant with the plant-word). When this happens the plant-word becomes part of the agent's permanent lexicon.

15.3.3. *Simulation / Population Dynamics*

Given a computational model of an individual agent we need to set out the ways in which a population of agents interacts. Before creating a population of agents, the environment is created containing different plants (which represent different meanings). The plants are distributed around the environment and they cover 2.28% of it.

The population in the simulations described here is made up of differing numbers of agents that are distributed randomly in the environment at the start. Agents form different communities each of whose members initially share a common lexicon. Agents can be either of the same or different type (CR or introspective) within the community. Apart from the differences in the initial lexicons and types between the agents, all other properties are the same.

Once the simulation starts the agents begin walking randomly in the environment. At every time step each agent moves to a random position, and looks for other agents (that fall into his field of vision). If he sees another agent then two of them will enter a dialogue where the 'seer' will be the speaker and the 'seen' the addressee. A dialogue is of the form:

- Agent 1 sees Agent 2 → **speaker** sees **addressee**.
- **speaker** looks around himself for a topic of conversation (a plant). If:
 - no plants in the field of vision, **speaker** sends *goodbye* to **addressee** and both of them walk off. Otherwise:
 - if plants in vision, **speaker** chooses a random plant as a topic for conversation.
- **speaker** sends the representation string for the chosen plant to the **addressee**.
- **addressee** tries to ground the plant-word via lexicon look up. There is no attempt to verify that the perceived meaning is same as the intended meaning.

- if **addressee** has the plant-word in her lexicon, the dialogue is considered *successful* and both agents continue walking.
- else if the **addressee** doesn't know the plant-word she resorts to CR or introspection (depending on her type) in order to acquire the meaning of the plant-word.

After the completion of the dialogue the agent continues walking in a random direction. This is then repeated until all agents acquire all the lexicon.

15.4. Results

This section describes different setups and experiment results for the model described in Sect. 15.3.3.

In order to test the questions raised in (7) the agents need to have minor but random lexical differences (here missing words), and clarification requesting (CR) and introspective capabilities.

The performance is based upon two behaviours which are collected at the end of the simulation run:

- *acquisition time*: the average time it takes the whole population to learn the whole lexicon.
- *convergence rate*: the percentage of *correctly* acquired meanings. Here an acquired meaning is *correct* if it is identical with the meaning associated with that word by the community who uses it at the start of the simulation. Note though that for one type of simulation, that involving a homogenous population of CR agents, this parameter has an entirely predictable value—100%. This is due to the fact that in the current set up each time an agent A makes a CR the original speaker explicitly provides A with the intended meaning associated with the unknown word. Thus at the end of such a simulation all the agents share a common lexicon with no divergence.

The initial conditions and model parameters affect the above behaviours in complex ways. To determine what consequences arise when a single parameter is manipulated there is a need to control all other parameters and keep them constant whilst only manipulating the parameter being investigated.

The three model parameters which we initially vary are:

- *meaning space*: the number of different meanings (plant-types) in the simulation.
- *population size*: the number of agents in the simulation.
- *acquisition threshold*: the number of times an agent has to hear an unknown word before she can acquire it. There are two types of acquisition threshold depending on the agent type:

- *CR threshold*: the number of times an agent has to hear an unknown plant-word before she can clarify it. For example if `CR-thres = 2` then an agent will only resort to a CR after hearing an unknown word for the second time.
- *Introspection threshold*: the number of times an agent has to associate an unknown plant-word with a plant-type in her field of vision before she can acquire the plant-word. For example if `Intr-thres = 4` then for every unknown plant-word that the agent hears, she has to see (associate) a plat-type four times with it before she is able to acquire the plant-word.

Each parameter has a default value throughout the experiments, unless it is being investigated. The default and investigative parameter values are shown in Table 15.1.

There are three types of experiments that we run where the different model parameters are tested. In the initial experiments (Sect. 15.4.1) the population is homogenous, either completely composed of CR agents or of introspective agents. In the second set of experiments (Sect. 15.4.2) the population is mixed, containing both CR and introspective agents in a 1:1 ratio. The final set of experiments (Sect. 15.4.3) is made up of populations of hybrid agents that can both ask a clarification request or introspect.

15.4.1. Homogenous Population Experiments

This section looks in detail at experiments with homogenous populations as described above. The results of these experiments serve as a benchmark for more complex population types. Manipulating the model parameters changes the behaviour of the simulation, where as stated above the performance of the populations will be judged on their *acquisition time* and *convergence rate*.

Meaning Space. In this set of experiments the parameter which is being manipulated is the meaning space. Increasing the meaning space involves increasing the differentiation among types of plants. The actual number of tokens remains constant. Increasing the meaning space causes the average acquisition time to go up polynomially (see Figure 15.1). This is the case for both CR and introspective populations, with CR performing slightly better. The reason for this is that an introspective agent has to associate an unknown

TABLE 15.1. Default and investigative parameter values.

| Parameter | Default value | Investigative values |
|-----------------------|---------------|------------------------|
| meaning space | 5 | 3, 4, 5, 10, 20 |
| population size | 40 | 4 – 40 |
| acquisition threshold | 3 | 1 – 10, 20, 30, 40, 50 |

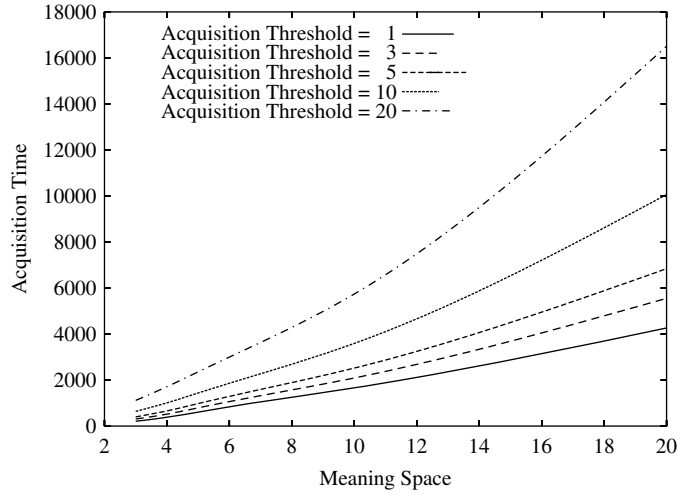


FIGURE 15.1. As the *meaning space* increases the *acquisition time* increases polynomially. Increasing the *acquisition threshold* causes the overall time to go up.

plant-word with the plants in her field of vision.² Sometimes it might happen that the addressee agent doesn't see any plants, or that she is looking at a different plant from the speaking agent, upon hearing an unknown plant-word. This increases the acquisition time, compared with a CR agent, who is given the meaning as soon as she requests clarification.

The convergence rate is always perfect in homogenous CR populations, thus we can only talk about the convergence rate in introspective populations for this set of experiments. The general trend here is that as the meaning space increases the convergence rate increases as well between 5 – 30% (depending on the inference threshold value) in the introspective population. The convergence rate increases for a meaning space up to a value of ten, then levels off afterwards (see Figure 15.2).

Population Size. The population size in these experiments increases from four agents up to forty agents whilst keeping the environment constant. The acquisition time decreases as the number of agents increases (see Figure 15.3). An explanation for this would appear to be that as more and more agents are placed in an environment of a constant size, the probability of seeing other agents at the next time step increases thus the probability of engaging in a dialogue also increases. Thus the time is dependant on the number of agents according to a power function $time = K * numAgents^{-a}$, where K and a are constants. This again holds for both CR and introspective populations. Here again the CR populations acquire their lexicons faster than the introspective populations.

² An agent's field of vision consists of a grid of fixed size originating from her location. Hence proximate agents have overlapping but not identical fields of vision.

15. Lexical Acquisition with/without Metacommunication 297

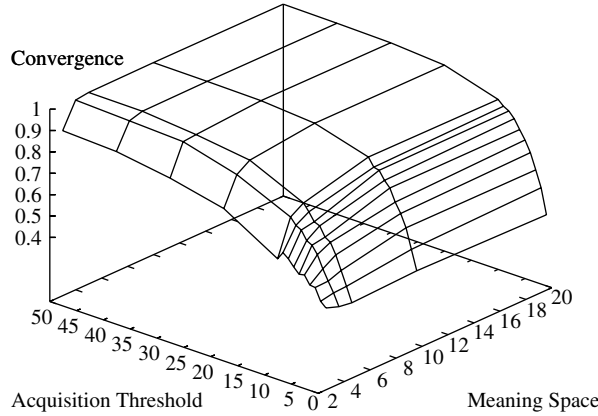


FIGURE 15.2. The relationship between *acquisition threshold*, *meaning space* and *convergence rate*: increasing the *acquisition threshold* causes the *convergence rate* to increase more than would be the case of increasing just the *meaning space*.

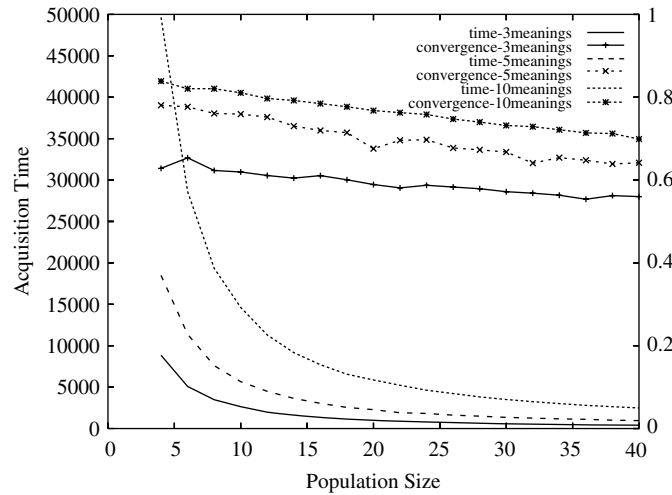


FIGURE 15.3. The relationship between *population size*, *acquisition time* and *convergence rate*. The smooth lines show the effect of increasing the *population size* on the *acquisition time* of the simulation. On the other hand the dotted lines represent the relationship between *population size* and *convergence rate*.

The convergence rate on the other hand decreases as the population size increases. The decrease in convergence rate is around 14% between a population size of four and forty. The reason for this might be termed the *Babel effect*. As the population increases the agents have more dialogues in a smaller part of the environment. So when an agent acquires a different meaning for a plant-word the agents around her are more likely to acquire this alternative meaning for an unknown plant-word causing the overall convergence rate to go down.

Acquisition Threshold. In these experiments the parameter that is changing is the number of times that an agent needs to hear a word before she is able to acquire it. For both types of populations the acquisition time increases linearly with the increasing acquisition threshold. As with the previous experiments the introspective populations are slower in acquiring the lexicon.

As for the convergence rate, it increases as the acquisition threshold increases. The reason is that an agent is given more chances to associate an unknown plant-word with the plants she sees as the threshold increases. Therefore she has more time to learn by observation, which improves the convergence rate. After a certain point the convergence rate starts to level off and even increasing the threshold values causes minimal change in the convergence rate. This can be seen in Figure 15.2.

15.4.2. *Mixed Population Experiments*

Now that we know how the parameters affect the simulation for homogeneous populations, we want to compare these results with the results of mixed populations of agents, containing both CR and introspective agents in a 1:1 ratio. In so doing, we want to keep the meaning space and the population constant and monitor how the manipulation of the acquisition thresholds affects the simulations. For the following experiments, the values for the meaning space and the population size used are as shown in Table 15.1.

In the CR community, the convergence rate increases from 90% to nearly 100%. On the other hand in the introspective community the convergence rate increases from 40% to 80%. This is shown in Figure 15.4(a). Increasing the CR threshold doesn't affect the convergence rate in any particular way as is shown in Figure 15.4(b). The convergence rate is much more dependent on the introspection threshold.

Comparing this result with the homogenous introspective population, the convergence rate is slightly better as shown in Figure 15.6. When averaging the overall convergence rate of both CR and introspective communities the convergence rate rises by quite a bit compared with the homogenous introspective population (see Figure 15.6).

The acquisition time doesn't seem to be affected for the differing population makeups. Only when calculating the overall average time (for both CR and introspective agents), do we see an improvement comparing it with the homogenous introspective population, but it is still higher than the homogenous CR populations (see Figure 15.7).

15.4.3. *Hybrid Agent Experiments*

In this final set of experiments the population is homogeneous, every agent has a capability of either using the CR strategy or the introspective strategy. Upon

15. Lexical Acquisition with/without Metacommunication 299

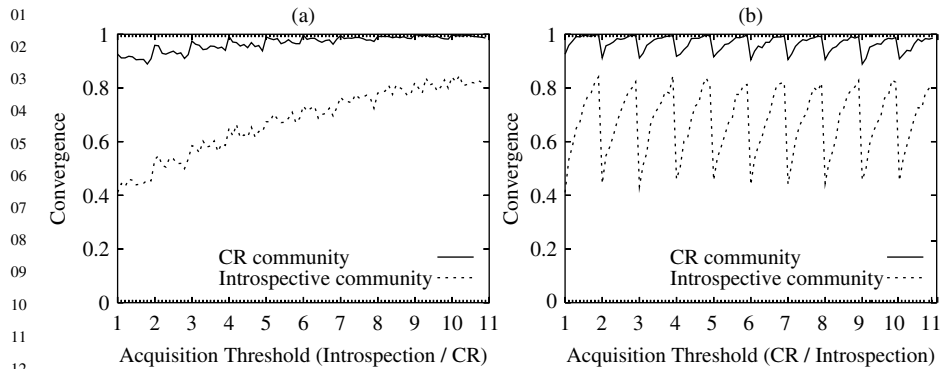


FIGURE 15.4. (a) A mixed population made up of a community of CR and a community of introspective agents (each containing 20 members) with differing *acquisition thresholds*. The x-axis shows different *acquisition threshold* make-ups, where the bigger line represents the introspection threshold and within each value there are 10 different CR threshold values (represented by the small lines on the x-axis). The upper curve represents the *convergence rate* of the CR community, while the lower curve represents the *convergence rate* of the introspective community. (b) This figure shows the same data as Fig. 15.4(a) but here the x-axis plots differing introspection threshold (small lines on the x-axis) values for the increasing CR threshold values.

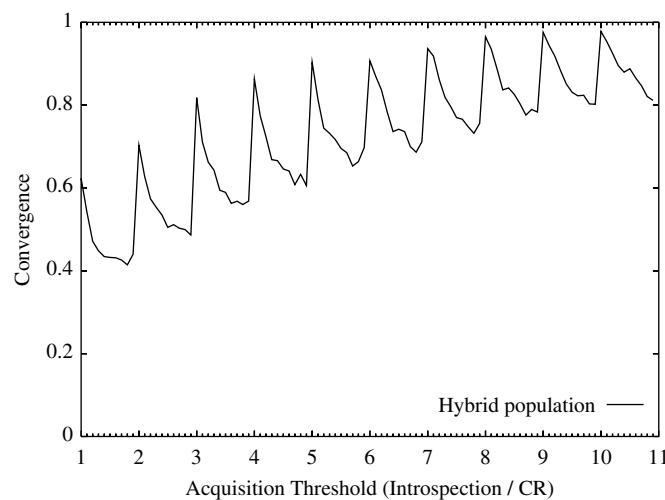


FIGURE 15.5. Hybrid population results: *convergence rate* is plotted against differing values for the *acquisition thresholds*, where for each introspection threshold value (represented on the x-axis with bigger gaps) CR threshold is being increased from 1 to 10 (represented by the smaller lines).

hearing an unknown plant-word the agent looks for plants close by. If she can see some plants, then she follows the introspective strategy, otherwise if there are no plants in her field of vision she resorts to a clarification request.

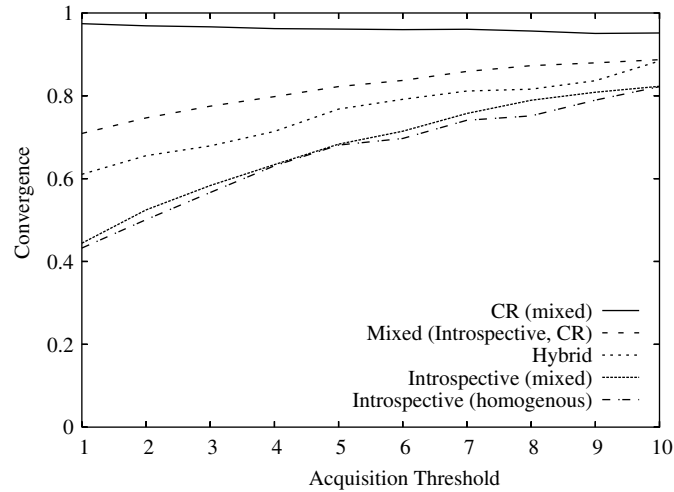


FIGURE 15.6. *Convergence rate increases as the acquisition threshold increases for every population make-up except for CR (mixed) population for which it decreases slightly.*

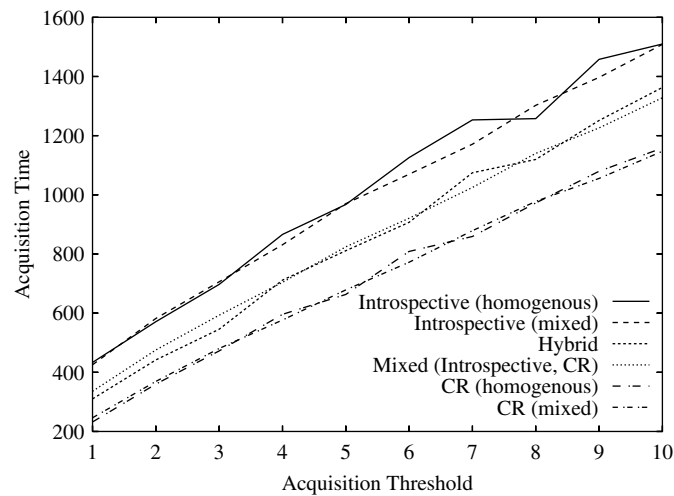


FIGURE 15.7. *The comparison of acquisition times for the experiments on differing population make-ups. Changing the population from homogenous to mixed (either introspective or CR) doesn't affect the acquisition time. The overall acquisition time for a mixed population (average of Introspective (mixed) and CR (mixed)) is comparable with the acquisition time of the Hybrid population.*

The results of these simulations are shown in Figure 15.5. The graph shows that as the introspection threshold increases so does the convergence rate. But the convergence rate also depends on the CR threshold, and as the CR threshold rises so the convergence rate falls. The reason for this is that as CR threshold rises,

the agent will resort to clarification requests less frequently thus introspecting more often, a strategy which is more error prone. If the CR threshold value is low and the introspection threshold value high then the convergence rate outperforms all the other tested populations. The overall average convergence rate is still higher than the homogenous and mixed introspective populations (see Figure 15.6).

Timewise the hybrid population is comparable with the overall mixed population (as shown by Figure 15.7).

15.5. Discussion

Let us now sharpen our focus to concentrate on the key results that emerge from these simulations. Possibly the most fundamental result obtained here concerns the convergence rate of a homogeneous population of introspective agents. In particular, the notable convergence decrease as a population *increases* seems to point to an important phenomenon, which we dubbed the *Babel effect*. The Babel effect means that a population without CRs is in danger of arriving at a situation where apparently a single language is shared by the population, whereas in fact divergent meanings are associated with the same sounds.³ Moreover, this dangerous situation could emerge in relatively quick time given that acquisition time decreases rapidly as the number of agents increase.

The counterpart to the Babel effect are the results concerning a population of hybrid agents, agents who use both introspection and CRs (the latter sort of as a last resort): such a population (as long as its members do not act recklessly, i.e. jump too quickly to conclusions about meanings) supports high rates of convergence.

15.6. Conclusions and Future Work

In this paper we have discussed how metacommunicative interaction (MCI) serves as a key component in the maintenance of a linguistic interaction system. We have outlined the basic components that need to emerge in order that an MCI-less linguistic system evolves into an MCI-containing system. This theoretical setting underpins and motivates the development of an ALife environment in which the lexicon dynamics of populations that possess and lack MCI capabilities are compared. The environment is one in which agents walk randomly and when proximate to one another engage in a brief conversational interaction concerning plants visible to the agents. We ran a series of experiments whose initial state involved agents possessing distinct lexicons and whose end state was

³ Though there is some evidence that perhaps this is not so unusual a situation, see (Schober 2004).

one in which all agents associated meanings with each word in a lexicon. The experiments involved tracking two variables: acquisition time and convergence rate—the percentage of newly acquired words whose acquired meanings match the originally associated meanings. Several parameters were varied (size of meaning space, population size, acquisition threshold—the number of times an agent has to hear an unknown word before she can acquire it.) and results gathered across distinct types of populations: the essential contrast being between agents who can request clarification of unknown words as opposed to agents who acquire new meanings introspectively, by observing their environment. The overall results can be viewed in Figure 15.7 and Figure 15.6. The main effect demonstrated, one we dub the *Babel effect*, is that the convergence rate of population that relies exclusively on introspection is intrinsically bounded and, moreover, this bound decreases with an increasing population. This bound seems to disappear once agents are endowed with CRs.

While the Babel effect seems to be an interesting finding, confirming our initial theorizing, much work remains to buttress it as a fundamental dividing line between MCI-ful and MCI-less populations. A significant simplification inherent in our simulation is the nature of linguistic interaction, which in our case involves syntactically unstructured messages. This intrinsically restricts the size and variation of the meaning space. An additional simplification is the built in success of responses to CRs. More open ended issues, whose role the simulation brings out, revolve around the influence of topography on lexicon dynamics (e.g. variety of plant types in the environment, ease of interaction between agents). An important future development to our simulation, obviously crucial for any evolutionary model, concerns mortality: as things stand, (for a single simulation run) agents are immortal and convergence rates are measured once all agents assign a meaning to each linguistic form. Limiting life span should reduce convergence and raise issues of generational variation.

Acknowledgments This work is supported by an EPSRC quota award to Macura and by grant RES-000-23-0065 from the Economic and Social Council of the United Kingdom to Ginzburg. We would like to thank two reviewers for EELC05 for their very helpful comments.

References

- Bates, E.: The Emergence of Symbols. Academic Press (1979)
- Briscoe, E. J.: Grammatical acquisition: Inductive bias and coevolution of language and the language acquisition device. *Language*, **76** (2000) (2):245–296
- Clark, H.: Using Language. Cambridge University Press, Cambridge (1996)
- Collier, N., Howe, T., and North, M.: Onward and Upward: The Transition to Repast 2.0, Proceedings of the First Annual North American Association for Computational Social and Organizational Science Conference, Electronic Proceedings, Pittsburgh, PA USA June (2003)
- Ginzburg, J.: The emergence of metacommunicative meaning. Paper presented at the 13th Amsterdam Colloquium on Semantics and Logic, December (2001)

15. Lexical Acquisition with/without Metacommunication 303

- Ginzburg, J., Cooper R.: Clarification, ellipsis, and the nature of contextual updates. *Linguistics and Philosophy*, **27** (2004) (3):297–366
- Ginzburg, J., Sag, I. A.: Interrogative Investigations: the form, meaning and use of English Interrogatives. Number 123 in CSLI Lecture Notes. CSLI Publications, Stanford: California (2000)
- Greenfield, P., Rumbaugh, S. S.: Comparing communicative competence in child and chimp: The pragmatics of repetition. *Journal of Child Language*, **20** (1993) 1–26
- Grice, H. P.: *Studies in the Ways of Words*. Reprinted from a 1957 article. Harvard University Press (1989)
- Kirby, S.: Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge University Press (2000) 303–323
- Ninio, A., Snow, C.: *Pragmatic Development*. Westview (1996)
- Poesio, M., Traum, D.: Conversational actions and discourse situations. *Computational Intelligence*, **13** (3) (1997)
- Repast. Recursive porous agent simulation toolkit. <http://repast.sourceforge.net>
- Schegloff, E.: Repair after the next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, **97** (1992) (5):1295–1345
- Schober, M.: Just how aligned are interlocutors representations? Commentary on Pickering and Garrod. *Behavioral and Brain Sciences*, **27** (2004) 209–210
- Steels, L.: The origins of ontologies and communication conventions in multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, **1** (1998) (2):169–194